ABSTRACT OF THE DISCLOSURE

An object such as a book or the like which contains text is photographed as a moving image, a still image is extracted from the moving image and undergoes broad-range identification to identify a text region, and image information in the text region is converted into text information, thus generating document data which can be processed later.